ETSI White Paper No. #52

ETSI Activities in the field of Artificial Intelligence

Preparing the implementation of the European AI Act

1st Edition – December -2022

ISBN No. 979108262073

Authors: Markus Mueck (Intel Deutschland GmbH), Raymond Forbes (Huawei Tech. (UK) Co. Ltd), Scott Cadzow (Cadzow Communications), Suno Wood (Association eG4U), Evangelos Gazis (HUAWEI TECH. GmbH), Christophe Gossard (John Deere GmbH & Co. KG), Francois Ortolan (NEC Europe Ltd), Lindsay Frost (NEC Europe Ltd), Hamed Farhadi (Ericsson), Matthias Schneider (Usability Labs), Martin Böcker (Human Factors)

VISIT…

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org

About the authors

Markus Mueck

INTEL Deutschland GmbH, Germany, Email: Markus.Dominik.Mueck@intel.com

Raymond Forbes

Huawei Tech. (UK) Co. Ltd, Email: raymond.forbes@huawei.com

Scott Cadzow

Cadzow Communications, Email: scott@cadzow.com

Suno Wood

Association eG4U, Email: sunowood@mac.com

Evangelos Gazis

HUAWEI TECH. GmbH, Email: vangelis.gazis@huawei.com

Christophe Gossard

John Deere GmbH & Co. KG, Email: GossardChristophe@JohnDeere.com

Francois Ortolan

NEC Europe Ltd, Email: Francois.Ortolan@EMEA.NEC.COM

Lindsay Frost

NEC Europe Ltd, Email: Lindsay.Frost@neclab.eu

Hamed Farhadi

Ericsson Research, Email: hamed.farhadi@ericsson.com

Matthias Schneider

Usability Labs, Email: msch@usability-labs.de

Martin Böcker

Human Factors, Email: boecker@humanfactors.de

Contents

About the authors 3

Contents 4

1 Executive Summary 5

2 ETSI's history of involvement in AI 7

3 Overview of world-wide regulation actions on Artificial Intelligence 8

3.1 Summary of (selected) global policy activities related to Artificial Intelligence 8

4 Europe: Societal Challenges in AI 10

4.1 Overview of Societal Challenges 10

4.2 European AI Act Requirements to address Societal Challenges 11

5 ETSI activities of relevance to address Societal Challenges in AI 12

5.1 Existing ETSI Deliverables of relevance to the AI Act 12

5.2 Planned ETSI Activities in support of the implementation of the AI Act 15

5.2.1 Human Factors requirements related to transparency, information to the users, and human oversight of AI systems 16

5.2.2 Glossary and Standards Landscape 20

5.2.2 AI Robustness, Trust & Confidence Building in AI-powered systems, and Test & Certification of selected class(es) of AI-powered systems 20

5.2.3 Testing-based conformity assessment for AI-enabled systems 21

5.2.4 Cognitive Management of AI Systems 22

6 Summary of ETSI Technical Committees and Industry Specification Group developing deliverables with relevance to the AI Act 22

7 Involvement of the European Research Community 25

8 Next Steps and Conclusion 26

Annex 1: Comparison European AI Act and US Blueprint for an AI Bill of Rights 27

Bibliography 28

1 Executive Summary

The European Parliament is currently preparing a major regulation in the field of Artificial Intelligence: The European Artificial Intelligence Act (AI Act) which is currently available as a public draft (Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, April 2021, available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 [1]) and as a recent compromise revision (Interinstitutional File 2021/0106(COD), 11 November 2022, available at https://artificialintelligenceact.eu/wp-content/uploads/2022/11/AIA-CZ-Draft-General-Approach-11-Nov-22.pdf [23]), which has significant updates likely to be approved, including listing of critical digital infrastructure as a high-risk use case for AI Systems in Annex III of the regulation.

The present White Paper provides information to concerned stakeholders, including SMEs, Industry, Academia, Government Regulation Agencies and others, on the current implementation status of standards potentially suitable for ensuring compliance to the original draft of the AI Act, from an ETSI perspective. The overall set-up within ETSI is discussed and most relevant Technical Committees and Industry Specification Groups and related available deliverables and plans are identified.

The implementation of the AI Act is expected to involve all three European Standards Organizations (ESOs) with the objective to develop supporting specifications, including Harmonised European Norms (HENs). HENs are typically used by manufacturers for the purpose of demonstrating compliance with the essential requirements of the regulation with the benefit of speeding product development and reducing the burden on statutory testing authorities. The impact of the AI Act will extend far beyond the borders of the EU to address the global market supply to the EU, and the EU will set the standard for management of AI for the global stage. ETSI, as a leading player on that global stage, will seek to spearhead that global reach.

In the present White Paper, ETSI offers a summary of societal challenges outlined in the AI Act and details activities in its Technical Committees and Industry Specification Groups which are of relevance to the AI Act and thus can be exploited and driven forward for its implementation. ETSI has conducted a survey across its technical activities and has mapped its technical activities to specific (sub-)Articles of the AI Act that are presented in.

Through participation in ETSI of all stakeholders involved in the standardization process in support of the AI Act, direct influence can be made on the definition of specific technical requirements and testing procedures in support of the AI Act. Stakeholders thus have the unique opportunity to shape the related framework and ensure its suitability to maintain continued product and service access to the European Single Market.

Measurement and evaluation of an AI system does require a few fundamentals to be in place to achieve trust in AI and this White Paper addresses the role of ETSI in enabling them:

- Metrics: metrics, benchmarks and thresholds (if applicable).
- Test: Test procedures to test the models and improve them if need be.
- Evaluation: An evaluation of the final model can then be used to “certify” a certain level of robustness/security.

Trust of itself is difficult to define, but this White Paper contends that by ensuring the fundamental building blocks are in place, trust can be built into the role and application of AI.

The application of many of the provisions of the AI Act should not be undertaken carelessly or in such a manner that adherence to regulation impedes the provision of value to the user and community. For example, understanding and explaining the behaviour of an AI-based system controlling the electric charging of a vehicle is less important than for a system that autonomously drives the same vehicle!

2 ETSI's history of involvement in AI

ETSI has a long and active history in the development of Artificial Intelligence (AI) and systems that use and support AI. The term AI has been poorly or too widely defined in many cases and this has meant that often the term has been avoided in favour of adopting more focused terms: Zero-touch network and Service Management (in ISG ZSM) dealing with autonomous network management; Experiential Networked Intelligence (in ISG ENI) dealing with autonomous learning to optimize network performance and architectures. AI is a vast field involving a multitude of distinct expertise where, often, AI is not the end goal but a means to achieve the goal. For this reason, ETSI has chosen to implement a distributed approach to AI – specialized communities meet in technically focused Technical Committees (TCs) and Industry Specification Groups (ISGs). Examples include TC Cyber with a specific focus on Cybersecurity aspects, ISG SAI working towards securing AI systems, ISG ENI dealing with the question of how to integrate AI into a network architecture, etc.

The role of AI as a means of enhanced machine-enabled decision making has in fact been at the heart of many groups in ETSI for a number of years. For example, in eHealth the first edition of the use case document published in 2019, already acknowledged the role of machine based proxies for health professionals, and this has become increasingly more nuanced in subsequent editions. The concerns raised around security have been addressed since September 2019 in the ISG SAI. The autonomous management of networks has been explored for some time before the creation of ISG ZSM with the publication of a White Paper in October 2016 describing GANA (the Generic Autonomic Networking Architecture), and the White Paper from October 2017 that set the foundation for ISG ENI, and then to the White Paper "Artificial Intelligence and future directions for ETSI" from June 2020 that can be considered as a forefather of the current document.

3 Overview of world-wide regulation actions on Artificial Intelligence

3.1 Summary of (selected) global policy activities related to Artificial Intelligence

There are several policy initiatives ongoing in various regions across the world related to regulating Artificial Intelligence Systems. A summary of selected key initiatives is given in the present section.

- European Artificial Intelligence Act (AI Act) [1]

  The European Commission, the European Parliament and the European Council are jointly preparing the European Artificial Intelligence Act (AI Act) with a first draft being available [1] and with revisions under development [23]. The objectives of this initiative are stated as follows:

  1. ensure that AI systems placed on the European Union market and used are safe and respect existing law on fundamental rights and European Union values,
  2. ensure legal certainty to facilitate investment and innovation in AI,
  3. enhance governance and effective enforcement of existing law on fundamental rights (e.g., GDPR (ETSI TR 103 747 V1.1.1 (2021-11) [5]) and safety requirements applicable to AI systems,
  4. facilitate the development of a single market for lawful, safe and trustworthy AI applications and prevent market fragmentation.

  Concerning the definition of an AI System, the draft AI Regulation complements the definitions of the Organisation for Economic Co-operation and Development with at least one of the three main paradigms of 'intelligence' (see Annex I of ETSI TR 103 473 V1.1.2 (2018-12) [6]): *A software that is developed with one or more of the techniques and approaches listed in Annex I of [6]) and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*. Note that this definition is still under debate in the EU Parliament and Council and may differ in the finally published AI Act.

- US Blueprint for an AI Bill of Rights (BLUEPRINT FOR AN AI BILL OF RIGHTS MAKING AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE, OCTOBER 2022, US White House, available at https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf [2])

  The *Blueprint for an AI Bill of Rights* [2] is a set of five principles and associated practices to help guide the design, use, and deployment of automated systems to protect the rights of the American public in the age of Artificial Intelligence. Developed through extensive consultation with the American public, these principles are a blueprint for building and deploying automated systems that are aligned with democratic values and to protect civil rights, civil liberties, and privacy.

The five principles are as follows:

- Safe and effective systems,
- Algorithmic discrimination protection,
- Data privacy,
- Notice and explanation,
- Human alternatives, consideration and fallback.

- US NIST AI Risk Management Framework (AI RMF) (see AI Risk Management Framework: Second Draft, August 2022, National Institute of Standards and Technology (NIST), available at https://www.nist.gov/itl/ai-risk-management-framework [3]).

  The NIST AI Risk Management Framework (AI RMF) is intended for voluntary use and to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems.

  The Framework is being developed through a consensus-driven, open, transparent, and collaborative process that includes workshops and other opportunities to provide input. It is intended to build on, align with, and support AI risk management efforts by others. A second draft of the AI RMF is currently available for feedback.

- The UK pro-innovation approach to regulating AI (UK Government policy paper on establishing a pro-innovation approach to regulating AI, UK Government, July 2022 [13]).

  The UK is proposing to establish a pro-innovation framework for regulating AI which is underpinned by a set of cross-sectoral principles tailored to the specific characteristics of AI:

  - Context-specific,
  - Pro-innovation and risk-based,
  - Coherent,
  - Proportionate and adaptable.

  A roadmap to an effective AI assurance ecosystem is detailing the steps required to build AI assurance ecosystem in the UK.

- The Japanese guidelines on AI Governance for Implementation of AI principles (METI) [29]

  This Guideline provides practical guidance for AI system operators as well as AI system developers.

While some initiatives are ongoing across multiple regions, it is desirable to achieve some level of consistency. Ideally, market access requirements are globally aligned with the objective to avoid distinct product designs adapted to potentially diverging requirements for each region.

4 Europe: Societal Challenges in AI

4.1 Overview of Societal Challenges

One of the key objectives of the EU AI Act is to address AI related Societal Challenges in Europe. A summary of key societal challenges is given below as defined in [1]:

- Ensure fundamental rights of persons
    - As stated by [1], AI should be a tool for people and be a force for good in society with the ultimate aim of increasing human well-being. Rules for AI available in the Union market or otherwise affecting people in the Union should therefore be human centric, so that people can trust that the technology is used in a way that is safe and compliant with the law, including the respect of fundamental rights.
- Ensure health of persons
    - For example, in health, the European health data space will facilitate non-discriminatory access to health data and the training of Artificial Intelligence algorithms on those datasets, in a privacy-preserving, secure, timely, transparent and trustworthy manner, and with an appropriate institutional governance.
- Ensure safety of persons
    - While safety risks of AI systems ensuring safety functions in machinery are addressed by the requirements of this Regulation, certain specific requirements in the [Machinery Regulation] will ensure the safe integration of the AI system into the overall machinery, so as not to compromise the safety of the machinery as a whole.

Separate draft legislation regarding legal liability for violation of the above principles (see ISO 25119 [24]) recognizes, however, that current national liability rules, based on fault, are not suited to handling liability claims for damage caused by AI-enabled products and services. Under such rules, victims need to prove a wrongful action or omission by a person who caused the damage, which may be difficult or prohibitively expensive for victims.

As it will be further outlined in the next section, ETSI hosts a community of experts who are able to address all the upper societal challenges as well as additional objectives which have been added to the 4th compromise proposal of the European AI Act as published on 19th October 2022 [15]:

- ensure that AI systems placed on the market or put into service in the Union are safe and respect Union values and strengthen the Union's digital sovereignty open strategic autonomy,
- promote investment and innovation in AI, including through increasing legal certainty, as well as competitiveness and growth of the Union market,
- enhance multistakeholder governance, representative of all relevant European stakeholders (e.g., industry, SMEs, civil society, researchers),
- contribute to strengthening global cooperation on standardisation in the field of AI that is consistent with Union values and interests.

Furthermore, separate draft legislation regarding legal liability for violation of the above principles (see COM (2022) 496: Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to Artificial Intelligence (AI Liability Directive [25] recognizes, however, that current national liability rules, in particular based on fault, are not suited to handling liability claims for damage caused by AI-enabled products and services. Under such rules, victims need to prove a wrongful action or omission by a person who caused the damage, which may be difficult or prohibitively expensive for victims.

4.2 European AI Act Requirements to address Societal Challenges

The European AI Act [1] introduces a number of Articles of different nature. Some of the key technical requirements towards an AI System are summarized in Table 1.

Table 1: Article of EU AI Act introducing technical requirements.

Requirements	Summary as defined by the European AI Act [1] (see AI Act for full details)
Data and data governance	High-risk AI systems … shall be developed on the basis of training, validation and testing data sets that meet the quality criteria ...
Technical documentation	The technical documentation shall be drawn up in such a way to demonstrate that the high-risk AI system complies with the requirements …
Record keeping	High-risk AI systems shall be designed and developed with capabilities enabling the automatic recording of events (‘logs’) …
Transparency and information to users	High-risk AI systems shall … ensure that their operation is sufficiently transparent to enable users to interpret the system’s output and use it appropriately …
Human oversight	High-risk AI systems shall be designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use …
Accuracy robustness and cybersecurity	High-risk AI systems shall … achieve, in the light of their intended purpose, an appropriate level of accuracy…
Risk management system	A risk management system shall be established, implemented, documented and maintained in relation to high-risk AI systems …

Quality management system	Providers of high-risk AI systems shall put a quality management system in place that ensures compliance with this Regulation …

It is expected that a Standardisation Request (SR) may be issued to European Standardisation Organizations (ESOs) to develop standards in support of the above requirements. It is thus essential to involve relevant expert communities to develop suitable standards, including Harmonised European Norms, to support manufacturers to provide compliant products and services to the Single European Market.

5 ETSI activities of relevance to address Societal Challenges in AI

As a general observation across available AI related deliverables published by international SDOs, very few standards actively address societal challenges and make mandates that apply to society; however, all standards, and the products and services they relate to, are to a greater or lesser extent influenced by the society they will be deployed in and on behalf of. In ETSI's TC eHEALTH, for example, this is particularly noted with regards to identifying requirements to ensure that devices and services which support diagnostic and therapeutic healthcare are held to a higher level of accountability to ensure that the societal expectation of ethics and patient care are visible. This level of expectation is also addressed in bodies that contribute to the frameworks used in eHealth such as the work in ISG SAI on explicability and transparency of AI processing, and in the work of SmartM2M, ISG CIM and others, on assurance of clear semantic and contextual labelling of data, backed by assurances made in TC CYBER of security-by-default, and privacy-by-design. Taken as a whole the explicit message of SDOs is that they take societal challenges seriously by working toward common societal goals of safe, secure and accountable use of data and services across all activity.

5.1 Existing ETSI Deliverables of relevance to the AI Act

As introduced above, ETSI has taken a distributed approach towards AI with specialized communities meeting in focused Technical Committees (TCs) and Industry Specification Groups (ISGs). In order to have an accurate picture of its own activities, ETSI organized a dialogue with its various communities and has identified a set of ETSI deliverables which are of direct relevance to the AI Act and are recommended to be exploited for its implementation. The identified list of available ETSI deliverables relevant to support the implementation of the AI Act is summarized below:

- **ETSI Deliverables of relevance to Draft AI Act "Article 9 – Risk Management System"**
    - ETSI TS 103 195-2 V1.1.1 (2018-05) [4]: Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management
    - ETSI TR 103 748 V1.1.1 (2022-06) [18]: Core Network and Interoperability Testing (INT); Artificial Intelligence (AI) in Test Systems and Testing of AI Models; Use and Benefits of AI Technologies in Testing

- Draft TR 103 749 [21]: INT Artificial Intelligence (AI) in Test Systems and Testing AI models; Testing of AI with definition of quality metrics
- ETSI TS 102 165-1 [30]: Part 1: Method and pro forma for Threat, Vulnerability, Risk Analysis (TVRA) (A revision is in development to further address AI as a threat agent and as a countermeasure accelerator)
- ETSI ISG SAI GR-004 [31]: Problem Statement
- ETSI ISG SAI GR 001 [32]: Ontology
- Draft ETSI GR SAI 009 [33]: Artificial Intelligence Computing Platform Security Framework
- ETSI GR SAI 005 [34]: Mitigation Strategy Report

- **ETSI Deliverables of relevance to Draft AI Act "Article 10 – Data and Data Governance"**
  - ETSI TR 103 747 V1.1.1 (2021-11) [5]: Core Network and Interoperability Testing (INT/ WG AFI); Federated GANA Knowledge Planes (KPs) for Multi-Domain Autonomic Management & Control (AMC) of Slices in the NGMN(R) 5G End-to-End Architecture Framework
  - ETSI TS 103 195-2 V1.1.1 (2018-05) [4]: Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management
  - ETSI TR 103 473 V1.1.2 (2018-12) [6}: Evolution of management towards Autonomic Future Internet (AFI); Autonomicity and Self-Management in the Broadband Forum (BBF) Architectures
  - ETSI TR 103 404 V1.1.1 (2016-10) [7]: Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Autonomicity and Self-Management in the Backhaul and Core network parts of the 3GPP Architecture
  - ETSI TR 103 626 V1.1.1 (2020-02) [8]: Autonomic network engineering for the self-Future Internet (AFI); An Instantiation and Implementation of the Generic Autonomic Network Architecture (GANA) Model onto Heterogeneous Wireless Access Technologies using Cognitive Algorithms
  - ETSI TR 103 627 V1.1.1 (2022-05) [9]: Core Network and Interoperability Testing (INT/WG AFI) Autonomicity and Self-Management in IMS architecture
  - ETSI GS CIM 009 V1.6.1 (2022-08) [10]: cross-cutting Context Information Management (CIM); NGSI-LD API
  - ETSI GS ENI 005 V2.1.1 (2021-12) [11]: Experiential Networked Intelligence (ENI); System Architecture
  - ETSI GR ENI 009 V1.1.1 (2021-06) [17]: Experiential Networked Intelligence (ENI); Definition of data processing mechanisms
  - Draft ETSI EG 203 922 [35]: The role of AI in eHealth.

- **ETSI Deliverables of relevance to Draft AI Act "Article 11 – Technical Documentation"**
    - Work is available within ETSI TC INT.
- **ETSI Deliverables of relevance to Draft AI Act "Article 12 – Record Keeping"**
    - ETSI GS ENI 005 V2.1.1 (2021-12) [11]: Experiential Networked Intelligence (ENI); System Architecture
    - ETSI GS PDL 011 V2.1.1 (2022-09) [18]: Permissioned Distributed Ledger (PDL); Specification of Requirements for Smart Contracts' architecture and security
    - ETSI GR PDL 014 V1.1.1 (2022-10) [12]: Permissioned Distributed Ledger (PDL); Study on non-repudiation techniques
    - ETSI TS 103 195-2 V1.1.1 (2018-05) [4]: Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management
    - ETSI TR 103 473 V1.1.2 (2018-12) [6]: Evolution of management towards Autonomic Future Internet (AFI); Autonomicity and Self-Management in the Broadband Forum (BBF) Architectures
    - ETSI TR 103 404 V1.1.1 (2016-10) [7]: Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Autonomicity and Self-Management in the Backhaul and Core network parts of the 3GPP Architecture
    - ETSI TR 103 626 V1.1.1 (2020-02) [8]: Autonomic network engineering for the self-Future Internet (AFI); An Instantiation and Implementation of the Generic Autonomic Network Architecture (GANA) Model onto Heterogeneous Wireless Access Technologies using Cognitive Algorithms
    - ETSI TR 103 627 V1.1.1 (2022-05) [9]: Core Network and Interoperability Testing (INT/WG AFI) Autonomicity and Self-Management in IMS architecture
    - Work is available within ETSI TC INT.
- **ETSI Deliverables of relevance to Draft AI Act "Article 13 – Transparency and Provision of Information to Users"**
    - ETSI GS ENI 005 V2.1.1 (2021-12) [11]: Experiential Networked Intelligence (ENI); System Architecture
    - ETSI ISG SAI GR 007 [36]: Explicability and transparency of AI processing
    - ETSI ISG SAI GR 010 [37]: Traceability of AI models
    - Work is available within ETSI TC INT.

- **ETSI Deliverables of relevance to Draft AI Act "Article 14 – Human Oversight"**
    - ETSI GS ENI 005 V2.1.1 (2021-12) [11]: Experiential Networked Intelligence (ENI); System Architecture
    - Draft ETSI TR 103 629 [19]: Evolution of Management towards Autonomic Future Internet (AFI); Confidence in autonomic functions; Guidelines for design and testability
- **ETSI Deliverables of relevance to Draft AI Act "Article 15 – Accuracy, Robustness and Cybersecurity"**
    - ETSI GR PDL 014 V1.1.1 (2022-10) [12]: Permissioned Distributed Ledger (PDL); Study on non-repudiation techniques
    - Draft TR 103 857 [20]: Autonomic Management and Control (AMC) Intelligence for Self-Managed Fixed & Mobile Integrated Networks (AFI); Generic Framework for E2E Federated GANA Knowledge Planes for AI-powered Closed-Loop Self-Adaptive Security Management & Control, Across Multiple 5G Network Slices, Segments, Services and Administrative Domains
    - Draft GR SAI 009 [38]: Artificial Intelligence Computing Platform Security Framework
    - Work is available within ETSI TC INT.
- **ETSI Deliverables of relevance to Draft AI Act "Article 17 – Quality Management System"**
    - ETSI TR 103 748 V1.1.1 (2022-06) [16]: Core Network and Interoperability Testing (INT); Artificial Intelligence (AI) in Test Systems and Testing of AI Models; Use and Benefits of AI Technologies in Testing
    - Draft TR 103 749 [21]: INT Artificial Intelligence (AI) in Test Systems and Testing AI models; Testing of AI with definition of quality metrics

It is further recognized across ETSI's output that from the simplest of editing rules and work preparation that standards should be designed in support of the fundamental rights of persons. This is reflected in the use of a gender-neutral language, of clear justifications for any mandate, and of clear accountability for the output of ETSI.

5.2 Planned ETSI Activities in support of the implementation of the AI Act

It is noted that the extent to which existing standards can help to underpin the high-level AI Act requirements vary dramatically for different areas. For more process-oriented requirements which largely rely on organizational or general technical measures (e.g., documentation, quality management), a lot of things can probably be reused, while for the core technical aspects (e.g., robustness, cybersecurity) there are no silver bullets yet and the topics are still intensely being researched.

Many ETSI Technical Committees and Industry Specification Groups are currently planning, or actively engaged in, specific activities in support of the implementation of the AI Act, in addition to the wider role and application of AI in systems. Below, a summary is given on the current plans.

5.2.1 Human Factors requirements related to transparency, information to the users, and human oversight of AI systems

This activity addresses user-related aspects of AI systems. They deal with (a) transparency of AI use (how users request and are provided with information and how transparent the reasoning underlying the provided information is) and (b) the interaction of humans with AI systems (human oversight).

Transparency and Information to the User

The following user-related issues of the use of AI-based systems affect the usability (in terms of effectiveness, efficiency, and satisfaction of use) as well as the accessibility of those systems in terms of being useful for and usable by users with the widest range of capabilities with a focus on transparency and information provided to the users.

In this document, “users” is a generic term for various roles, from technical expert in charge of system performance, to a citizen chosen randomly from a population with a wide range of physical and mental capabilities: therefore, the specific requirements must later be tailored to the actual roles and not just the generic case.

Use cases include humans

- as users of public information systems (e.g., public warning systems),
- as users of consumer products (e.g., smart homes),
- as users of professional equipment (both in the private and public sector), and
- as members of Human-Machine collaborative systems.

Issues that are candidates for standardisation in this area concern explainability (or explicability), transparency, expectations of the users, expectations of the AI system, trust, and ways of ensuring accessibility.

Explainability:

Systems using Machine Learning that are based on Artificial Neural Networks in most cases cannot explain how they arrived at a result. This is due to the nature of neural networks. For example, a Machine-Learning based system may point to an area of a picture that it classified as containing a road junction, but it cannot explain the defining characteristics of a road junction that it detected in that picture. This is recognition (e.g., detected by image correlation to a set of learned images of road junctions) as opposed to deductive reasoning (e.g., noting that several roads are likely to intersect hence the point of intersection being a junction).

For human users, this is a novel situation given that they are used to receiving recommendations from other humans who can relate to their reasoning (regardless of whether the advice is sound or not).

Wherever possible, AI systems should offer a rationale for their reasoning. Depending on the side conditions, the target audience (e.g., lay user or expert) and the criticality of the application, such rationales may take different forms. Starting from relatively high-level information (e.g., derived from model cards), rationales may be refined up to sophisticated analyses based on XAI methods. Since providing a sufficient rationale will not always be possible, users need, therefore, be trained on the

limitations of some AI systems regarding their capabilities of explaining and justifying results. In other words, using those systems requires trust.

Transparency:

Some rules may be designed into the system, or arise as unexpected consequences, without the user's knowledge.

For example, future health insurance policies may make insurance premiums dependent on the users' behaviours, without the customers being aware of those rules. If the system detects that the user stopped smoking, his premium may go down. If, on the other hand, it detects that his 15-year-old daughter started dating a boyfriend, rates may go up again as she may become pregnant.

Users should be made aware of the existence of built-in rules in AI systems that impact the users' well-being, be that financial, social, medical, etc.

Furthermore, AI systems will have checks in place to help detect potential rules or trends (biases) which arise during operations.

Expectations of the Users:

Users who have no accurate mental model of a system's capabilities may overestimate what a system can do at a given time. A well-known example of such an overestimation of an AI-based system's capabilities are the reports of users in cars with advanced levels of automation who have overestimated, or over interpreted, the capability, e.g., drivers who elected to sleep during a trip expecting the car in "autopilot" mode to operate at maximum automation levels while in fact the system only provided partial automation (SAE level 2).

The most important approach to make sure that users' expectations match the actual capabilities of the system is to communicate the prerequisites and limitations of the system to the users in an unambiguous, transparent and understandable way.

Expectations of the System:

The capabilities of an AI-based system may be limited in such a way that it expects user intervention when it encounters situations it cannot handle alone. Users then should be prepared to intervene. An example is autonomous car control at a low or medium automation level. In cars of that type, users may encounter situations in which they are handed over of car controls with an advanced warning of a certain time span (typically, a few seconds). This means users will take over the controls (steering wheel, brake pedal, etc.) and establish situation awareness (understanding what is currently happening, what is going to happen, and what they should do). Sleep or alertness detection may be used by the system to make sure users are ready to take control if required. The same technology can also be used for control room operators to ensure they are ready to take over control when required.

Trust:

The quality of an AI solution, including those that employ deep learning techniques, depends inter alia on the quantity and quality of the data it has been trained with. The user needs to know about the reliability of the system before trusting it.

Trusting an AI-based system has several important preconditions:

- The AI system needs to be able to explain the quality of its reasoning, at least in terms of the extent of the data with which it has been trained. This means that any limited reliability or trustworthiness of the results is indicated to the user.
- The user needs to be able to question the reasoning of a system and the system will be able to justify, or set within understandable bounds, its conclusions or actions.
- If the users identify decisions which might indicate a certain bias (e.g., an AI-based recruiting system that favours males), which cannot be explained by the system, they will have the option to request an audit of the AI-system to ensure that the system does not create solutions to the detriment of certain user groups.
- The activity for which the AI-system is employed cannot not harm third parties. This puts certain requirements on the application areas in which AI systems can or may be employed. Understanding and explaining the behaviour of an AI-based system controlling the electric charging of a vehicle is less important than for a system that autonomously drives the same vehicle!

Accessibility:

This area deals with the requirements of people with special needs or disabilities when using AI-based systems. Examples of these uses are AI-based appliances or mobility aids in smart homes owned and inhabited by elderly and/or disabled people. Other examples include smart traffic solutions in cities, which are used by people with specific needs, or smart energy meters employed by energy providers to optimize power consumption.

For systems deployed in these usage scenarios, requirements should be defined to ensure that people with accessibility needs can successfully interact with and control these systems.

Human Oversight of AI Systems

The following user-related issues of the use of AI-based systems affect the usability (in terms of effectiveness, efficiency, and satisfaction of use) as well as the accessibility of those systems in terms of being useful for and usable by users with the widest range of capabilities with a focus on the human oversight of such systems.

Use cases include humans working with AI systems

- as operators (in the widest sense), or
- as members of Human-Machine collaborative systems.

The following topics need extensive analysis which will lead to usability-related requirements for human controlling and or collaborating with AI-Systems.

Requirements on human control of AI systems cover the whole lifecycle of system use:

- information prior to acquisition,
- initial setup,
- personalisation and preferences (including touch, voice, gesture and other interaction styles),
- intended use,
- prevention of use that the system is not intended for,
- feedback,
- abort of misunderstood commands,
- interpretation of information and/or system states resulting from a user command,

- system updates, and
- safe decommissioning.

The extent to which human-in-the-loop contributions to the operation of an AI system is required depends on the maturity of the technology for offering safe, effective, and efficient operation for achieving a result. This depends inter alia on the level of automation that is supported (e.g., form “0 No automation” to “5 Full automation”, with no human intervention required). Humans will often be involved to express intentions (e.g., to invoke a smart-home functionality or to start a machine operation). The degree of human involvement will also be subject to a trade-off balancing human risk and equipment damage (e.g., machines searching for and neutralising land mines). This also includes residual risk, as defined e.g., in ISO 25119 [24].

A further set of requirements affects humans working together with other humans and AI systems. Humans cooperating on specific tasks have well-developed means of controlling and steering the necessary communication between them. If an AI-based system becomes part of such a cooperating system, it is to be ensured that the communication between the AI-based system and human partners/controllers follows the rules of human-human cooperation. To ensure such a successful communication between cooperating parties, there needs to be means to ensure that all cooperating parties (humans and machines) have the same basic knowledge and share the same basic understanding of the objectives of the system and the cooperation. This includes:

- the ability to create common beliefs: team members (Human and AI) have shared beliefs about the world state, the goals, the conventions associated with the task at hand, etc., all on a jointly accepted level of abstraction,
- the capability of all parties to predicting each other’s actions with a reasonable degree of accuracy,
- the capability to redirect each other’s behaviour (controllability),
- a common understanding of the degree of automation offered by the machine component (e.g., farmers work with harvesting tractors that are fully automated vs those that may require human intervention in certain situation – see human in the loop) and
- training measures for safe, efficient, effective, and satisfactory human-machine collaboration.

To ensure this cooperation the AI system follows certain rules of human-human cooperation. These requirements are all based on usability and accessibility needs of the human parties in such a cooperation.

The second aspect of such cooperating human-machine-system training measures for safe, efficient, and effective human-machine-cooperation which allow the human parties to adapt to restrictions and requirements of the machine (AI) component. For these training measures requirements are to be established to ensure that the training is adapted to the needs of users.

Human Factors Requirements and Test Procedures

The result of the proposed work are Human Factors requirements on the usability and accessibility of AI systems and a set of test procedures allowing manufacturers to document compliance with those requirements.

5.2.2 Glossary and Standards Landscape

Any new standardisation work on Artificial Intelligence should be conducted in the context of (a) existing standards and ongoing standardisation work, and (b) a shared conceptual basis, i.e., a glossary of relevant terms to be used in the documents to be developed.

Standards landscape

There are many international and European standards dealing with properties of software systems which are related to the requirements set out in the call for proposals namely to develop the standards framework for the following:

- risk management system,
- governance and quality of datasets used to build AI systems,
- record keeping through built-in logging capabilities,
- transparency and information to users,
- human oversight,
- accuracy specifications,
- robustness specifications,
- cybersecurity specifications,
- quality management system for providers, including post-market monitoring processes,
- conformity assessment.

To make sure that work already performed for non-AI-based ICT systems is not duplicated for AI-based systems, a landscaping activity identifying existing standards which might be re-used unmodified or revised for AI-based systems is urgently needed. One task will be to identify and/or develop this standards landscape. Existing overviews from EC JRC, CEN, ISO, StandICT, IEEE et al. should of course be referenced.

Glossary of relevant terms

A shared terminology of agreed-upon terms (words and expressions) is a prerequisite for the successful development of standards documents on AI. This is even more important as AI is a fast-developing discipline and new approaches are constantly being proposed by both academia and industry. To ensure that everybody (standards writers and users) is “on the same page”, the activity will furthermore include a task related to the development of a glossary of terms to be used throughout the upcoming work phases.

5.2.2 AI Robustness, Trust & Confidence Building in AI-powered systems, and Test & Certification of selected class(es) of AI-powered systems

The work on robustness, confidence building, testing, and certification has the following targets:

- The objective for this work is to develop methods and standards for achieving robustness in AI-based systems through appropriate test methods that enable the AI-based (AI-powered) systems to be improved in quality and robustness—thanks to the test methods, with consideration of a selected scope of class or classes of AI-powered systems or networks (e.g., the class of Autonomic or Autonomous Networks (ANs) that are powered by AI/ML.

- To develop methods that can be used for trust and to confidence building in AI-powered systems, with consideration of a selected scope of class or classes of AI-powered systems (e.g., the class of Autonomic or Autonomous Networks (ANs) that are powered by AI/ML. The ATS concept also applies in this aspect.

- To develop standardizable quality metrics for use as basis for test and certification of AI/ML Models for selected class or classes of AI powered systems or Networks.

- To develop standards to regulate the basic security requirements on AI computing platform which is regarded as the infrastructure of AI/ML system, aiming to guarantee a solid foundation for the system, and to be used as basis for test and certification of AI/ML system infrastructure.

- To develop a Guide that can be used by the industry towards development of enablers for testing and certification of selected class or classes of AI-powered systems, including the idea of Building an Ecosystem for Certification Labs for AI models and AI systems.

5.2.3 Testing-based conformity assessment for AI-enabled systems

With regard to "Testing-based conformity assessment for AI-enabled systems", the rationale of this activity is based on the fact that providers of 'high-risk' AI systems will be required to:

- Set up a risk management system and a post-market monitoring system,
- draw up technical documentation, which can be used as the basis for assessing conformity with the technical requirements in the regulation,
- put a quality management system in place that include "techniques, procedures and systematic actions for design, design verification, development, quality control, quality assurance" as well as "examination, test and validation procedures".

Conformity assessment is conducted by a third party or based on internal control; the main subject of the assessment is a technical documentation of the AI system.

In this context, requirements for quality characteristics for systems based on AI can be elaborated with a view to trustworthiness, including safety and security, explainability, transparency, reliability, accuracy, maintainability, fairness, degree of autonomy and controllability by humans and accountability. As part of test specifications, requirements for different quality criteria can be formulated for AI-based systems. The requirements can, e.g., be made dependent on the characteristics of data-driven and rule-based systems, including the description of relevant quality characteristics of ML-based components and the identification of specific risks related to ML-based components in AI-enabled systems. A corresponding documentation scheme that supports the continuous and consistent documentation of quality and quality-related attributes for AI-enabled systems is therefore essential for comparable and high-value documentation.

As failures due to different root causes, e.g., algorithm, implementation, training data, hyperparameters may have different implications with regard to the risk and conformity assessment, it is important to pursue a differentiated approach for the assessment, taking into account root-cause and cause-effect analysis with corresponding methods for the identification, classification, and documentation of causes for failures and their implications.

The scope of the first part of the activity is pre-standardization activity towards a catalogue of unified guidelines, interfaces and procedures based on a structured risk assessment and risk classification scheme.

5.2.4 Cognitive Management of AI Systems

Concerning Cognitive Management, existing reports and studies will be further extended to develop technical requirements, related implementation solutions and approaches to testing. The proposed activity will specifically address the following items:

- Explicability and transparency of AI processing,
- Privacy aspects of AI/ML systems,
- Traceability of AI models,
- Security testing of AI,
- Cognitive Network Management architecture of AI,
- Transformer Architecture and Policy translation in language processing including AI,
- Policy models in AI,
- Mitigation of Bias,
- Human-Centredness,
- Ethics,
- Trustability (which requires the above items along with other features).

The Specification of Policy and Cognitive Management enables the data processing and normalisation against known criteria, this making AI work in the way that communication manager expects, without surprises and within the law. Therefore, the Standardization of the Policy and Cognition is considered essential.

6 Summary of ETSI Technical Committees and Industry Specification Group developing deliverables with relevance to the AI Act

Beyond the deliverables summarized above, the following ETSI technical bodies are currently developing further guides, analysis (as ETSI Technical Reports) and standards that aim to serve as a starting point for the implementation of the AI Act with a specific focus on the upper Societal Challenges:

- ETSI ISG SAI (Securing Artificial Intelligence) has taken a global lead in considering the security implications of AI from an application-agnostic perspective, addressing the threats to and from AI. Starting from a thorough analysis of the threats to AI systems (ETSI GR SAI 001 [39] and ETSI GR SAI 004 [40]) it has gone on to study existing approaches for mitigating threats during the whole life cycle of AI systems. It considered measures aimed at the data supply chain (ETSI GR SAI 002 [41]) and countermeasures to a whole range of AI-specific attacks (ETSI ISG GR SAI 005 [42]). The documents yield useful input for technically underpinning the high-level requirements on security set out in the AI Act. SAI is also working on many other aspects including secure hardware and computing platforms, privacy, transparency and explicability as well as the misuse of AI for creating deepfakes.

- ETSI TC INT (Core Network and Interoperability Testing) has defined the Generic Autonomic Network Architecture (GANA) [see refs [4] – {9]) as an architectural reference model for autonomic networking, cognitive networking and self-management which can serve as a basis for the implementation of the AI Act.
- ETSI TC MTS provides technologies, tools, and guidelines on conformance and interoperability testing and certification of protocols and other systems, including AI and IoT systems, that are under standardization at various ETSI groups and committees. It interacts with other groups within ETSI and beyond in all matters related to testing and specification methodologies. The Testing and Test Control Notation version 3 (TTCN-3) is used for the standardized test suites at ETSI and 3GPP and is also published under ITU-T.
- ETSI ISG CIM (cross-cutting Context Information Management) defines a Context Information Management API (see ETSI GS CIM 009, V1.6.1 (2022-08) [10]) which allows users to provide, consume and subscribe to context information in multiple scenarios and involving multiple stakeholders. Context information is modelled as attributes of context entities, also referred to as "digital twins", representing real-world assets (e.g., a bus in a city or a luggage claim ticket). AI solutions are expected to consume data from such systems, and perhaps to use such knowledge graphs to monitor/report their own outputs.

- ETSI TC eHealth acknowledges a critical role for AI in the future provision of health services. The role of AI is also relevant to the areas of strategy for 2023, defined by the ETSI Board as “Socio-economic trends, Technology trends and Policy trends.” Health-related Use Cases for AI have stimulated public interest following the Covid-19 pandemic. Medical practitioners and service providers alike have noted the potential of AI for eHealth. In short, AI can support the work of public health authorities and governments to make effective policy decisions in the provision of diagnostics, crisis management, treatment, monitoring and control of long-term disease. TC eHEALTH has opened a new Work Item, DEG/eHEALTH-0016: The role of AI in eHEALTH. This work item will address the role of AI as an accelerator for eHealth processing. It will address the ethical dimension, the security dimension and the privacy dimension amongst others. The work should identify actions across the ETSI Technical Bodies to support more detailed future work.

- ETSI ISG ENI (Experiential Networked Intelligence) introduces an experiential architecture (see ETSI GS ENI 005, V2.1.1 (2021-12) [11]) (i.e., an architecture that uses Artificial Intelligence (AI) and other mechanisms to improve its understanding of the environment, and hence the operator experience, over time).

- ETSI ISG ZSM (Zero-touch network and Service Management) defines the required end-to-end architecture and solutions for network automation. The ultimate automation target is to enable largely autonomous networks to be driven by high-level policies and rules; these networks would be capable of self-configuration, self-monitoring, self-healing and self-optimization without further human intervention. All this requires a new horizontal and vertical end-to-end architecture framework designed for closed-loop automation and optimized for data-driven machine learning and Artificial Intelligence algorithms.

- ETSI ISG PDL (Permission Distributed Ledgers) covers the non-repudiation challenges in Permissioned Distributed Ledgers (PDLs), the non-repudiation strategies/technologies, and their viability in PDLs (ETSI GR PDL 014 V1.1.1 (2022-10) [12]). It also defines the limitations in non-

repudiation strategies in PDLs and possible future directions. ETSI ISG PDL GR 014 [12] furthermore discusses PDL based end-to-end architecture that provides non-repudiation. This includes nonrepudiation for input and output data for a PDL, such as external PDLs and smart contracts.

As a general observation, many ETSI TBs are directly addressing AI and the societal challenge. To give a specific example, TC eHEALTH, as previously mentioned, is active in ensuring that AI as a tool in health is subject to the same rigour that more conventional human based decision making is subject to. In the Intelligent Transport sector, represented by ETSI's TC ITS, the role of data in feeding the second-by-second decisions in traffic and vehicle management is very conscious of the impact of poor data and poor decision making has on the safety and well-being of all citizens, and is actively addressing the need for transparency and explicability in how AI, and wider data driven, decisions are arrived at. The purpose of these activities, in part, is to raise the confidence of wider society in the role that such AI and data driven activities do to address the societal concerns and challenges they directly or indirectly introduce.

7 Involvement of the European Research Community

The field of Artificial Intelligence is a top notch research area of key interest to the European Research Community. ETSI has established an ETSI Board Strategy group called “Research, Innovation and Standardisation Ecosystem (RISE)” which is actively working on an outreach to Europe’s Research community. The objective is to guide the available talent pool towards contributions in ETSI and exploit the available expertise to develop ETSI deliverables in support of the AI Act. ETSI Board RISE is in particular interacting with key platforms including NetworldEurope, the 6G Infrastructure Association (6G-IA) and others to maximize its outreach and to facilitate the access to ETSI or this community.

A specific collaboration is maintained with the European 6G Flagship project, called Hexa-X. Hexa-X is providing a vision and is developing key technological enabler for a future 6th generation cellular communications system. As a first result of the close collaboration between ETSI and Hexa-X, a new Industry Specification Group has been created with a focus on THz communication (ISG THz). ETSI and Hexa-X are further collaborating to exploit the findings and results of Hexa-X in the field of Artificial Intelligence and Machine Learning with the objective to support the implementation of the AI Act.

Furthermore, the European 6G Flagship project Hexa-X is providing a vision and is developing key technological enablers for a future 6th generation (6G) cellular communications system. It is envisioned for 6G to have a crucial role and responsibility for large-scale deployments of Artificial Intelligence in the wider society. 6G will provide a framework to support, enhance, and ultimately enable real-time trustworthy AI services – transforming AI/Machine Learning (ML) technologies into a vital and trusted tool for significantly improved efficiency and service experience. Hexa-X recognizes the necessity to expand the fundamental network design paradigm from mainly performance-oriented to both performance- and value-oriented. Here, value entails intangible yet important human and societal needs such as sustainability, trust, and inclusion. This will lead to a new class of evaluation criterion, i.e., Key Value Indicators (KVIs) that needs to be understood, developed, and adopted in the network design towards 6G. To fully embrace such a vision, Hexa-X has developed technical enablers (see Analysis of 6G architectural enablers applicability and initial technological solutions, October 2022 [26]) and architectural enablers (see AI-driven communication & computation co-design: initial solutions, June 2022 [27]) for AI-enabled 6G networks for improving network efficiency, preserving privacy, security, and trust. In addition, several key performance metrics (KPIs) and KVIs has been introduced for AI as a service in 6G networks and the target values are quantified for several AI-enabled 6G use cases (see Targets and requirements for 6G – initial E2E architecture, March 2022 [28].

Hexa-X is happy to align with EC and to support the EC in the implementation of European policy activities in the field of AI, e.g., by introducing the KPI, KVIs and the technical solutions related to the AI to be used in 6G and to the AI services to be enabled by 6G in line with the AI Act requirements.

8 Next Steps and Conclusion

ETSI is committed to supporting the European Commission in the implementation of the AI Act. As outlined in the present White Paper, a number of ETSI deliverables are available which may be used as a basis for working towards the implementation of the AI Act. ETSI hosts a number of complementary expert communities, each with a specific focus and competence. Those communities have expressed their commitment to be involved in the development of related ETSI deliverables. Corresponding planned activities are outlined in the present paper.

In case that a Standardisation Request (SR) is being issued, ETSI is available for a close collaboration with the European Commission to ensure a successful implementation of the AI Act. For this purpose, ETSI will build on its existing community of AI experts as well as ETSI's close ties to the European Research landscape.

Annex 1: Comparison European AI Act and US Blueprint for an AI Bill of Rights

A high level comparison of the EU AI Act [1] and the US Blueprint for an AI Bill of Rights [2] is outlined by Table **A.1**A.1. It is observed that there is alignment between at least some of the underlying principles and requirements. Such an alignment is indeed desirable by manufacturers in order to be able to apply consistent design principles to AI products across all regions.

Table A.1: High-Level comparison EU AI Act and US Blueprint for an AI Bill of Rights.

Requirement	EU AI Act	US Blueprint for AI Bill of Rights
Risk Management System	Article 9	Safe and effective systems
Data and data governance	Article 10	Algorithmic discrimination protection; Data privacy
Transparency and provision of information to users	Article 13	Notice and explanation
Human Oversight	Article 14	Human alternatives, consideration and fallback

Note that there is further interrelation between the EU AI Act [1] and the US Blueprint for an AI Bill of Rights [2] on the more detailed level of respective requirements.

Bibliography

[1] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, April 2021, available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

[2] BLUEPRINT FOR AN AI BILL OF RIGHTS MAKING AUTOMATED SYSTEMS WORK FOR THE AMERICAN PEOPLE, OCTOBER 2022, US White House, available at https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

[3] AI Risk Management Framework: Second Draft, August 2022, National Institute of Standards and Technology (NIST), available at https://www.nist.gov/itl/ai-risk-management-framework

[4] ETSI TS 103 195-2 V1.1.1 (2018-05): Autonomic network engineering for the self-managing Future Internet (AFI); Generic Autonomic Network Architecture; Part 2: An Architectural Reference Model for Autonomic Networking, Cognitive Networking and Self-Management

[5] ETSI TR 103 747 V1.1.1 (2021-11): Core Network and Interoperability Testing (INT/ WG AFI); Federated GANA Knowledge Planes (KPs) for Multi-Domain Autonomic Management & Control (AMC) of Slices in the NGMN(R) 5G End-to-End Architecture Framework

[6] ETSI TR 103 473 V1.1.2 (2018-12): Evolution of management towards Autonomic Future Internet (AFI); Autonomicity and Self-Management in the Broadband Forum (BBF) Architectures

[7] ETSI TR 103 404 V1.1.1 (2016-10): Network Technologies (NTECH); Autonomic network engineering for the self-managing Future Internet (AFI); Autonomicity and Self-Management in the Backhaul and Core network parts of the 3GPP Architecture

[8] ETSI TR 103 626 V1.1.1 (2020-02): Autonomic network engineering for the self-Future Internet (AFI); An Instantiation and Implementation of the Generic Autonomic Network Architecture (GANA) Model onto Heterogeneous Wireless Access Technologies using Cognitive Algorithms

[9] ETSI TR 103 627 V1.1.1 (2022-05): Core Network and Interoperability Testing (INT/WG AFI) Autonomicity and Self-Management in IMS architecture

[10] ETSI GS CIM 009 V1.6.1 (2022-08): cross-cutting Context Information Management (CIM); NGSI-LD API

[11] ETSI GS ENI 005 V2.1.1 (2021-12): Experiential Networked Intelligence (ENI); System Architecture

[12] ETSI GR PDL 014 V1.1.1 (2022-10): Permissioned Distributed Ledger (PDL); Study on non-repudiation techniques

[13] UK Government policy paper on establishing a pro-innovation approach to regulating AI, UK Government, July 2022, available at

https://www.gov.uk/government/publications/establishing-a-pro-innovation-approach-to-regulating-ai/establishing-a-pro-innovation-approach-to-regulating-ai-policy-statement

[14] UK The roadmap to an effective AI assurance ecosystem, UK Government, December 2021, available at https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem

[15] European AI Act – 4th proposal, European Commission, October 2022, available at https://artificialintelligenceact.eu/wp-content/uploads/2022/10/AIA-CZ-4th-Proposal-19-Oct-22.pdf

[16] ETSI TR 103 748 V1.1.1 (2022-06): Core Network and Interoperability Testing (INT); Artificial Intelligence (AI) in Test Systems and Testing of AI Models; Use and Benefits of AI Technologies in Testing

[17] ETSI GR ENI 009 V1.1.1 (2021-06): Experiential Networked Intelligence (ENI); Definition of data processing mechanisms

[18] ETSI GS PDL 011 V2.1.1 (2022-09): Permissioned Distributed Ledger (PDL); Specification of Requirements for Smart Contracts' architecture and security

[19] Draft ETSI TR 103 629: Evolution of Management towards Autonomic Future Internet (AFI); Confidence in autonomic functions; Guidelines for design and testability

[20] Draft TR 103 857: Autonomic Management and Control (AMC) Intelligence for Self-Managed Fixed & Mobile Integrated Networks (AFI); Generic Framework for E2E Federated GANA Knowledge Planes for AI-powered Closed-Loop Self-Adaptive Security Management & Control, Across Multiple 5G Network Slices, Segments, Services and Administrative Domains

[21] Draft TR 103 749: INT Artificial Intelligence (AI) in Test Systems and Testing AI models; Testing of AI with definition of quality metrics

[22] https://hexa-x.eu/

[23] Interinstitutional File 2021/0106(COD), 11 November 2022, available at https://artificialintelligenceact.eu/wp-content/uploads/2022/11/AIA-CZ-Draft-General-Approach-11-Nov-22.pdf

[24] ISO 25119 "Tractors and machinery for agriculture and forestry – Safety-related parts of control systems", https://www.iso.org/standard/80216.html

[25] COM (2022) 496: Proposal for a DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on adapting non-contractual civil liability rules to Artificial Intelligence (AI Liability Directive), available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0496

[26] Analysis of 6G architectural enablers applicability and initial technological solutions, October 2022.

[27] AI-driven communication & computation co-design: initial solutions, June 2022.

[28] Targets and requirements for 6G – initial E2E architecture, March 2022.

[29] Japan guidelines on AI Governance for Implementation of AI principles (METI)

[30] ETSI TS 102 165-1 [30]: Part 1: Method and pro forma for Threat, Vulnerability, Risk Analysis (TVRA) (A revision is in development to further address AI as a threat agent and as a countermeasure accelerator)

[31} ETSI ISG SAI GR-004: Problem Statement

[32] ETSI ISG SAI GR 001: Ontology

[33] Draft ETSI GR SAI 009: Artificial Intelligence Computing Platform Security Framework

[34] ETSI GR SAI 005: Mitigation Strategy Report

{35] Draft ETSI EG 203 922: The role of AI in eHealth.

[36] ETSI ISG SAI GR 007: Explicability and transparency of AI processing

[37] ETSI ISG SAI GR 010: Traceability of AI models

[38] Draft GR SAI 009: Artificial Intelligence Computing Platform Security Framework

[39] ETSI ISG SAI GR 001 V1.1.1 (2022-01): Securing Artificial Intelligence (SAI); AI Threat Ontology

[40] ETSI ISG SAI GR 004 V1.1.1 (2020-12): Securing Artificial Intelligence (SAI); Problem Statement

[41] ETSI ISG SAI GR 002 V1.1.1 (2021-08): Securing Artificial Intelligence (SAI); Data Supply Chain Security

[42] ETSI ISG SAI GR 005 V1.1.1 (2021-03): Securing Artificial Intelligence (SAI); Mitigation Strategy Report

ETSI
06921 Sophia Antipolis CEDEX, France
Tel +33 4 92 94 42 00
info@etsi.org
www.etsi.org